

Remarques sur la loi des erreurs.

Par CHARLES JORDAN à Budapest.

§ 1. Certaines questions concernant la loi des erreurs ne sont pas encore complètement élucidées. Il y a des personnes qui pensent que la loi des erreurs peut être démontrée, et qu'en l'appliquant il est possible de découvrir les erreurs dites systématiques, c'est-à-dire dues aux imperfections des instruments de mesure. Or il est évident qu'un fait d'expérience tel que "les erreurs d'observation suivent une certaine loi" ne peut être prouvé par aucune déduction mathématique. L'expérience seule peut en décider, mais là encore il y a des difficultés ; en effet, les faits d'expérience, justement à cause des erreurs d'observation inévitables, ne peuvent être conformes d'une manière absolue à aucune loi précise, ils ne peuvent la suivre que d'une manière approchée. Une autre difficulté consiste en ce que pour vérifier la loi il faudrait connaître la grandeur à mesurer, or cette grandeur est presque toujours inconnue. Dans l'exemple classique de la somme des angles d'un triangle, on mesure les *angles* et on montre que leur *somme* suit la loi ; la démonstration n'est pas directe. Pour la compléter il faut remarquer, (on le verra au § 4) que si l'on admet pour les erreurs la loi du hasard, il en résulte que la somme des erreurs la suit aussi. L'expérience fournira tout au plus la conclusion : Dans certaines conditions, les erreurs d'observation suivent d'assez près la loi dite du hasard.

Toutes les observations sont nécessairement affectées d'erreurs. En effet, si on mesure plusieurs fois successivement une même grandeur, les résultats seront différents et les différences seront dues aux erreurs d'observation. En général, on classe les erreurs en deux groupes, dans le premier on range les erreurs dites systématiques, c'est-à-dire produites par des causes con-

stantes, agissant toujours dans le même sens et causées, par exemple, par les imperfections des instruments ; l'observateur doit chercher à éliminer ces causes, car la théorie mathématique des erreurs ne peut rien dire sur ces erreurs.

Dans le second groupe, on classe les erreurs dites fortuites ou accidentelles, c'est-à-dire produites par de petites causes très nombreuses agissant tantôt dans un sens, tantôt dans un autre ; une telle erreur peut être considérée comme la résultante d'un très grand nombre de petites erreurs élémentaires n'ayant chacune que très peu d'influence sur le résultat. Une réflexion simple montre que ces dernières peuvent suivre une loi mathématique. En effet si l'on suppose par exemple que chaque cause élémentaire peut produire avec la même facilité une erreur $+1$ ou -1 , et si $2n$ causes sont en jeu, une erreur résultante 2ε peut être obtenue conformément au théorème de BERNOULLI de $\binom{2n}{n+\varepsilon}$ manières différentes et la probabilité de l'erreur 2ε sera :

$$\binom{2n}{n+\varepsilon} \frac{1}{2^{2n}}.$$

Lorsqu'on mesure une grandeur inconnue z et M_i est la mesure obtenue, l'erreur commise ε_i , inconnue aussi, sera $\varepsilon_i = M_i - z$. On supposera que cette erreur est accidentelle, et que la probabilité pour qu'elle soit comprise entre $\varepsilon_i - \frac{1}{2} \Delta \varepsilon_i$ et $\varepsilon_i + \frac{1}{2} \Delta \varepsilon_i$ est donnée par

$$(1) \quad P(\varepsilon_i) \Delta \varepsilon_i = \frac{1}{\sqrt{2\pi}} e^{-\varepsilon_i^2/2\sigma^2} \Delta \varepsilon_i;$$

c'est en même temps la probabilité pour que la mesure obtenue soit comprise entre $M_i - \frac{1}{2} \Delta M_i$ et $M_i + \frac{1}{2} \Delta M_i$, et on peut l'écrire :

$$P_1(M_i) \Delta M_i = \frac{1}{\sqrt{2\pi}} e^{-(M_i - z)^2/2\sigma^2} \Delta M_i.$$

C'est l'énoncé de la loi des erreurs. On en tire immédiatement la probabilité $\mathfrak{P}(\lambda)$ pour que la valeur absolue de l'erreur ε_i soit plus petite que λ ; on trouve

$$(2) \quad \mathfrak{P}(\lambda) = \frac{2}{\sqrt{2\pi}} \int_0^{\lambda/\sigma} e^{-\frac{1}{2}t^2} dt.$$

De la formule (1) on tire aussi celle de la probabilité des mesures simultanées M_1, M_2, \dots, M_n , c'est-à-dire celle des erreurs simultanées $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$

$$(3) \quad \mathfrak{S}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \Delta \varepsilon_1 \dots \Delta \varepsilon_n = \left(\frac{1}{\zeta \sqrt{2\pi}} \right)^n e^{-\sum \varepsilon_i^2 / 2\zeta^2} \Delta \varepsilon_1 \dots \Delta \varepsilon_n.$$

A l'aide de (1) on peut montrer que ζ^2 est l'espérance mathématique du carré des erreurs c'est-à-dire

$$\zeta^2 = \int_{-\infty}^{\infty} \varepsilon_i^2 P(\varepsilon_i) d\varepsilon_i.$$

Lorsqu'on fait n mesures, et que n est grand, ζ^2 est égal d'une manière approchée à la moyenne des carrés des erreurs

$$\zeta^2 \sim \frac{1}{n} \sum (M_i - z)^2;$$

ζ^2 est appelé "*dispersion des observations*" et sa racine carrée ζ "*erreur quadratique des observations*".

On peut considérer ζ comme la mesure de la précision des observations. En effet lorsque dans deux séries d'observations les erreurs quadratiques sont ζ_1 et ζ_2 , et qu'on trouve les probabilités égales pour que les erreurs soient respectivement plus petites que λ_1 et λ_2 , alors de (2) il résulte :

$$\frac{\lambda_1}{\zeta_1} = \frac{\lambda_2}{\zeta_2}$$

donc les grandeurs λ_1 et λ_2 correspondant à la même probabilité sont proportionnelles aux erreurs quadratiques ζ_1 et ζ_2 .

§ 2. Première remarque. La loi des erreurs n'est applicable qu'aux grandeurs M_i mesurées directement et non pas à une fonction quelconque $f(M_i)$ de ces mesures; autrement dit, l'erreur à considérer doit être la différence entre une mesure M_i et la grandeur à déterminer.

Il est vrai que certaines fonctions de M_i peuvent aussi suivre la loi des erreurs, mais cela doit être démontré dans chaque cas; de plus l'erreur quadratique de la fonction doit être déterminée en partant de celle des mesures. Cette remarque est utile, car il y a des auteurs qui appliquent la loi des erreurs à des fonctions de mesures sans s'inquiéter de la justifier.

Par exemple si l'on connaît les distances d_i d'un foyer de tremblement de terre aux observatoires O (pour $i=1, 2, \dots, n$)

et qu'on observe le temps t_i qu'une onde sismique met pour y arriver, et qu'on cherche à déterminer la vitesse v de l'onde, comme ce sont les temps qu'on a mesurés, les erreurs à considérer sont les grandeurs $\frac{d_i}{v} - t_i$ et non pas $\frac{d_i}{t_i} - v$ comme on l'a fait quelquefois.

Seconde remarque. La difficulté en employant la formule (1) est que les grandeurs z et ξ sont inconnues. Pour remédier à cet inconvénient, souvent on procède de la manière suivante : on fait n mesures et on détermine leur moyenne $\Sigma M_i/n$; puis on considère l'écart ξ_i entre la mesure M_i et la moyenne : $\xi_i = M_i - \frac{\Sigma M_i}{n}$, et finalement on remplace l'erreur ε_i par l'écart ξ_i , ce qui veut dire que l'on remplace z par la moyenne des mesures (on verra que c'est sa valeur la plus probable). Ce procédé ne conduit pas à des résultats satisfaisants. En effet si l'on désigne la moyenne des erreurs ε_i par η , alors de

$$\varepsilon_i = M_i - z$$

il suit

$$\eta = \frac{\Sigma M_i}{n} - z$$

donc η est aussi l'erreur de la moyenne des observations. De plus la définition de l'écart $\xi_i = M_i - \frac{\Sigma M_i}{n}$ conduit à $\Sigma \xi_i = 0$. Des équations précédentes on tire

$$(4) \quad \varepsilon_i = \xi_i + \eta$$

et par suite

$$\Sigma \varepsilon_i = n \eta.$$

En outre on a

$$\Sigma \varepsilon_i^2 = \Sigma \xi_i^2 + n \eta^2 = n(\sigma^2 + \eta^2)$$

ou, si n est grand

$$\zeta^2 \sim \sigma^2 + \eta^2.$$

GAUSS (*Werke*, Tome IV, p. 6) dit que la somme des erreurs accidentelles devrait être nulle, et si l'on trouve que la moyenne des erreurs est égale à η , alors η forme la *partie constante de l'erreur*, c'est-à-dire la partie produite par des causes agissant toujours dans le même sens. En retranchant η de chaque obser-

vation il obtient des "observations corrigées". Les erreurs correspondant à ces observations sont évidemment égales aux écarts ξ_i . Cette manière de voir a conduit ses disciples à des conceptions erronées, en faisant croire qu'il était possible de découvrir, à l'aide de la loi des erreurs, les erreurs constantes ou systématiques, par exemple celles dues aux défauts des instruments de mesure, ce qui est impossible. D'autre part on verra que l'erreur η de la moyenne des mesures est aussi une erreur fortuite qui suit la loi des erreurs.

De plus, en remplaçant dans ζ la grandeur z par sa valeur la plus probable, on trouve $\zeta^2 = \sigma^2$. Enfin en écrivant $\zeta = \sigma$ dans la formule (3) qui donne la probabilité simultanée des mesures M_1, \dots, M_n , on obtient

$$\left(\frac{1}{\sigma \sqrt{2\pi e}} \right)^n \Delta M_1 \dots \Delta M_n$$

ce qui conduit à des résultats inacceptables. On en conclut qu'on ne peut remplacer simplement dans les formules la grandeur ζ par sa valeur la plus probable; on verra qu'il est préférable de déterminer la probabilité totale de z quel que soit ζ .

Troisième remarque. Pour déterminer la probabilité à posteriori de la grandeur inconnue z , on est obligé d'avoir recours au théorème de BAYES. A cet effet il faut adopter une hypothèse concernant les probabilités à priori de z (probabilités avant les mesures); généralement on considère à priori toutes les valeurs de z comme également probables, c'est-à-dire on admet que la probabilité pour la grandeur inconnue d'être comprise entre $z - \frac{1}{2} \Delta z$ et $z + \frac{1}{2} \Delta z$ est proportionnelle à Δz . Cette hypothèse conduit à des conclusions très satisfaisantes. Il en résulte d'après le théorème de BAYES qu'après les n mesures, la probabilité à posteriori de z sera proportionnelle à (3), donc à

$$\left(\frac{1}{\zeta \sqrt{2\pi}} \right)^n e^{-\sum \varepsilon_i^2 / 2\zeta^2} \Delta \varepsilon_1 \dots \Delta \varepsilon_n \Delta z.$$

Il faut remarquer, que dans cette formule ζ est une fonction de z ; en effet on a, si n est grand, $\zeta^2 \sim \Sigma (M_i - z)^2 / n$; mais les mesures M_i dépendent aussi de z , de plus la relation précédente n'est pas valable si n est petit; elle ne peut servir par exemple à déterminer la dérivée de ζ par rapport à z . Pour sortir de cette

difficulté, le plus simple c'est de considérer la précision ζ des mesures dans une série d'observations indépendante des mesures M_i et de la grandeur z à mesurer. Comme ces quantités sont du même ordre de grandeur c'est très possible. Dans ces conditions on admettra que la probabilité à priori, c'est-à-dire avant les mesures, pour que la précision soit comprise entre $\zeta - \frac{1}{2} \Delta\zeta$

et $\zeta + \frac{1}{2} \Delta\zeta$ ne dépend que de ζ . A première vue on pourrait croire que toutes les valeurs de ζ sont également probables. Cette hypothèse conduirait aux résultats de GAUSS d'après lesquels la valeur la plus probable de ζ^2 est $\Sigma \xi_i^2 / (n-1)$; mais en d'autres cas elle donnerait des résultats moins satisfaisants. Du reste, l'idée que la probabilité d'une erreur quadratique très grande, par exemple plus grande que la grandeur à mesurer soit aussi probable qu'une erreur quadratique petite, est contraire à notre manière de penser. [D'après la formule (2), la probabilité pour que la valeur absolue de l'erreur ε_i soit plus grande que ζ est 0,32. Il en résulte que si n est assez grand, environ un tiers des mesures dépassera $z + \zeta$, ce qui est certainement impossible si $\zeta > z$].

On en conclut qu'il vaut mieux adopter l'hypothèse que la probabilité de ζ diminue quand ζ augmente, et de prendre cette probabilité proportionnelle à $\Delta\zeta/\zeta^\alpha$. Pour $\alpha=0$ cela donne l'hypothèse précédente. De plus on verra qu'il faut supposer α au moins égal à 2.

Quatrième remarque. GAUSS considère que la valeur la plus probable de ζ^2 est $n\sigma^2/(n-1)$, ce qui veut dire que la valeur la plus probable de la moyenne des carrés des erreurs est égale à la somme des carrés des écarts divisée par $n-1$, lorsqu'il s'agit d'une inconnue à déterminer, et divisé par $n-m$, lorsqu'il y en a m inconnues. (*Werke*, Tome IV, p. 49). L'idée est qu'étant données n équations et m inconnues ($n > m$) on peut considérer m équations comme satisfaites et $n-m$ équations donnant des erreurs; or même si l'on procédait ainsi, il faudrait tenir compte des erreurs nulles pour déterminer leur moyenne.

Cette remarque a une certaine importance, car la valeur adoptée pour ζ^2 influe sur les résultats numériques dans le calcul des moindres carrés. On verra que la valeur la plus probable de ζ^2 est $n\sigma^2/(n+\alpha-1)$; donc si on se contente d'une valeur

approchée, $\zeta^2 = \sigma^2$ est préférable à la valeur $\zeta^2 = n\sigma^2/(n-1)$ proposée par GAUSS.

Cinquième remarque. Quelquefois il faut faire entrer dans le calcul des mesures qui ne sont pas également précises. Pour y arriver il faut tenir compte de leur poids. On assigne à l'une des mesures, dont l'écart quadratique est ζ_1 , l'unité de poids et on dit qu'une seconde mesure dont l'écart quadratique est ζ_2 a un poids p , lorsqu'il faudrait combiner p observations de la première mesure pour que leur moyenne soit aussi exacte que la seconde mesure. Nous verrons au § 4 que l'erreur quadratique de cette moyenne est ζ_1/\sqrt{p} . On en conclut que les poids sont inversement proportionnels aux carrés des erreurs quadratiques :

$$\frac{p}{1} = \frac{\zeta_1^2}{\zeta_2^2}.$$

On verra au § 9, un exemple où l'on a déterminé le poids des valeurs x_i calculées, lorsque le poids des mesures M_i était considéré comme égale à l'unité.

On ne peut rien objecter à cette manière de procéder. Il arrive que les observateurs considèrent certaines mesures de la même grandeur comme plus exactes que les autres. Cette différence peut être due à ce que les mesures sont exécutées à l'aide d'appareils de précisions différentes, ou par des observateurs plus ou moins exercés, ou encore dans des circonstances différentes. Si dans ces cas, on désire assigner des poids différents aux mesures, il faut prendre des précautions, en tous cas il faut fixer les poids *avant l'exécution des observations*. Car si on les fixait après les mesures, comme cela a été fait plusieurs fois, par exemple en donnant à chaque mesure un poids d'autant plus petit que son écart de la moyenne est plus grand, on arriverait à des résultats en désaccord avec la loi des erreurs.

§ 3. La loi (1) des erreurs a été formulée pour la première fois par ADRAIN dans ses *Research concerning the probabilities of errors which happen in making observations* (*The Analyst or Mathematical Museum*, 1 (1808), pp. 93—109). Il en a tiré plusieurs applications, entre autres il a remarqué qu'en vertu de cette formule la valeur la plus probable d'une grandeur inconnue, mesurée plusieurs fois, est la moyenne arithmétique des valeurs observées; de plus que la position la plus probable d'un point déterminé dans l'espace est le centre de gravité des points observés.

Trois ans avant, LEGENDRE est arrivé à la même conclusion dans ces *Nouvelles méthodes pour la détermination des orbites des comètes* (Paris, 1806). Il dit: Étant donné la fonction

$$Y = a_1 x_1 + a_2 x_2 + \dots + a_m x_m$$

où les a_v sont des coefficients disponibles, on mesure le système de grandeurs x_v plusieurs fois; supposons que l' $i^{\text{ème}}$ mesure ait donné

$$Y_i = a_1 x_{1i} + a_2 x_{2i} + \dots + a_m x_{mi}$$

correspondant à une erreur $Y_i - Y$. Généralement le nombre des mesures est très grand (c'est toujours le cas dans les problèmes physiques et astronomiques), comme le nombre des équations surpasse celui des inconnues, il est impossible de disposer de ces inconnues pour annuler toutes les erreurs. Il faudra donc se contenter de les réduire autant que possible. Pour y arriver, différentes considérations ont amené LEGENDRE à disposer des inconnues de manière à rendre minimum la somme des carrés des erreurs $\Sigma(Y_i - Y)^2$. (C'est la méthode utilisée aujourd'hui.) Dans l'ouvrage cité ci-dessus il dit "La règle par laquelle on prend le milieu entre les résultats de différentes observations (p. 74) n'est qu'une conséquence très simple de notre méthode générale que nous appellerons *Méthode des moindres quarrés*". Plus loin il remarque que d'une manière plus générale, si plusieurs points sont donnés dans l'espace, la somme des carrés des distances de ces points à leur centre de gravité est minimum.

En 1809 GAUSS dans sa *Theoria Motus Corporum Coelestium* déduisait indépendamment de ce qui précède la même loi, et plus tard en 1823 et 1826 il a publié dans les *Commentationes Societatis Scientiarum Gottingensis* sa *Theoria Combinationis Observationum Erroribus Minimis Obnoxiae*. Dans cette dernière publication, il a donné une méthode complète de l'application de cette loi.

LAPLACE a déduit la loi des erreurs en partant de la supposition que chaque erreur est le résultat de la combinaison d'un très grand nombre d'erreurs élémentaires (*Théorie des probabilités*, 1812, p. 304). Plus tard YOUNG dans ses *Remarks on the probabilities of error in physical observations* (*Philosophical Transactions*, 1819) et HAGEN (*Grundzüge der Wahrscheinlichkeitsrechnung*, Berlin, 1837) ont perfectionné cette méthode. Plus tard d'autres

géomètres ont contribué à la théorie mathématique des erreurs, par exemple: CROFTON, TCHÉBICHEF, ENCKE, BIENAYMÉ, ELLIS, GLAISHER, LINDELÖF, STORY, CHARLIER, POINCARÉ.

Supposons que la grandeur z ait été mesurée n fois et que M_1, M_2, \dots, M_n soient les mesures obtenues. Désignons par $\varphi(\varepsilon_i)$ la probabilité de l'erreur $\varepsilon_i = x_i - z$. Si nous posons les conditions: 1. la probabilité de l'erreur ε est égale à la probabilité de l'erreur $-\varepsilon$, donc $\varphi(\varepsilon) = \varphi(-\varepsilon)$; 2. le maximum de $\prod_{i=1}^n \varphi(M_i - z)$ est atteint pour z égal à la moyenne des mesures M_i ; ces conditions conduisent à la loi (1).

Pour se rendre compte si certaines erreurs suivent la loi du hasard, remarquons que dans ce cas, si le nombre des mesures est grand, le nombre des erreurs positives doit être à peu près égal au nombre des erreurs négatives; la somme des erreurs doit être sensiblement égale à zéro. La moyenne des carrés des écarts divisée par le carré de la moyenne des valeurs absolues des écarts doit être à peu près $\frac{1}{2}\pi$.

Si la probabilité de l'erreur ε_i suit la loi (1), il faut que la probabilité totale de toutes les erreurs soit égale à l'unité. En réalité on a

$$\frac{1}{\sqrt{2\pi}} \int_{-z}^{\infty} e^{-\varepsilon^2/2z^2} d\varepsilon = \frac{1}{\sqrt{2\pi}} \int_{-z/\zeta}^{\infty} e^{-\frac{1}{2}t^2} dt \sim 1,$$

en effet lorsque z est grand relativement à ζ , cette intégrale est à peu près égale à l'unité. On en conclut que dans ce cas les erreurs peuvent, d'une manière approchée, suivre la loi (1) du hasard. Par contre lorsque les erreurs sont du même ordre de grandeur que la quantité z à mesurer, la loi n'est pas applicable.

§ 4. Pour obtenir la probabilité $\mathfrak{P}(\lambda)$ que la valeur absolue de l'erreur de la moyenne des n mesures soit plus petite que λ , c'est-à-dire $|\eta| < \lambda$, il faut effectuer la sommation:

$$\mathfrak{P}(\lambda) = \sum \Delta P(M_1, \dots, M_n)$$

pour toutes les valeurs de M_i satisfaisant à $|\sum \varepsilon_i| < n\lambda$.

Pour faciliter cette sommation considérons un facteur discontinu $F(n\lambda, \sum \varepsilon_i)$ tel que l'on ait $F=1$ lorsque $|\sum \varepsilon_i| < n\lambda$ et $F=0$ lorsque $|\sum \varepsilon_i| > n\lambda$. Comme la probabilité pour avoir exactement

$|\Sigma \varepsilon_v| = n\lambda$ est nulle, il suffit que F ait une valeur finie dans ce cas. Il en résulte qu'après avoir multiplié l'expression (3) par F , on peut faire la somme par rapport aux M_i de zéro à ∞ pour toutes les valeurs de i .

$$F = \frac{2}{\pi} \int_0^{\infty} \sin(n\lambda u) \cos(u \Sigma \varepsilon_v) \frac{du}{u}$$

est un tel facteur.

En multipliant (3) par F on trouve

$$\mathfrak{P}(\lambda) = \frac{2}{\pi} \left(\frac{1}{\zeta \sqrt{2\pi}} \right)^n \cdot \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\Sigma \varepsilon_v^2 / 2\zeta^2} \int_0^{\infty} \sin(n\lambda u) \cos(u \Sigma \varepsilon_v) \frac{du}{u} d\varepsilon_1 \dots d\varepsilon_n.$$

En remarquant que $\cos(u \Sigma \varepsilon_v)$ est la partie réelle de $e^{iu \Sigma \varepsilon_v}$, on aura des intégrales de la forme

$$\int_{-\infty}^{\infty} e^{-\varepsilon_v^2 / 2\zeta^2 + iu \varepsilon_v} d\varepsilon_v = \zeta \sqrt{2\pi} e^{-\frac{1}{2} u^2 \zeta^2}$$

pour toutes les valeurs de v . On en conclut

$$\mathfrak{P}(\lambda) = \frac{2}{\pi} \int_0^{\infty} e^{-\frac{1}{2} n u^2 \zeta^2} \sin(n\lambda u) \frac{du}{u}.$$

Posons encore, $\omega = \zeta u \sqrt{n}$ et $x = \lambda \sqrt{n} / \zeta$; la formule précédente deviendra

$$\mathfrak{P}(\lambda) = \frac{2}{\pi} \int_0^{\infty} e^{-\frac{1}{2} \omega^2} \sin(x\omega) \frac{d\omega}{\omega};$$

mais on sait que cette intégrale est égale à

$$(5) \quad \mathfrak{P}(\lambda) = 2 \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} t^2} dt$$

où $x = \lambda \sqrt{n} / \zeta$. On en conclut que $\mathfrak{P}(\lambda)$, la probabilité pour que la valeur absolue de l'erreur de la *moyenne* soit plus petite que λ est en même temps égale à la probabilité pour que la valeur absolue de l'erreur d'une *seule observation* soit plus petite que $\lambda \sqrt{n}$;

de plus elle est encore égale à la probabilité pour que la somme de erreurs de n observations soit plus petite que λn . En outre on en conclut, que l'erreur de la moyenne des mesures suit la même loi que l'erreur d'une simple mesure; de plus que si ζ est l'erreur quadratique des mesures M_i , alors celle de la moyenne sera égale à ζ/\sqrt{n} .

On en tire par exemple, que la probabilité est la même pour avoir

$$|\varepsilon_i| < 1 \quad \text{ou} \quad |\sum \varepsilon_i| < \sqrt{n} \quad \text{ou encore} \quad \left| \frac{\sum \varepsilon_i}{n} \right| < \frac{1}{\sqrt{n}}.$$

FOURIER connaissait cette règle bien avant que la théorie des erreurs fût découverte: "Pour prendre la hauteur de la pyramide de Chéops, il fit simplement mesurer par des soldats les 203 marches de ce gigantesque escalier. Vos hommes manquent d'habitude, disait-on, les surfaces sont irrégulières, les arêtes sont inclinées, aucune précision n'est possible, et l'erreur commise sur chaque marche sera multiplié par 203. Elle le sera par 14 seulement, répondit-il résolument car 14 est la racine carrée de 203." (BERTRAND, *Calcul des Probabilités*, p. XXXVIII.)

§ 5. Probabilité de l'erreur d'une fonction linéaire de m grandeurs mesurées. Soit

$$Y = a_1 z_1 + a_2 z_2 + \dots + a_m z_m.$$

Les coefficients a_ν sont donnés et on connaît les valeurs des z_ν ; supposons que les mesures des z_ν aient fourni les grandeurs $M_{\nu i}$; de manière que l' $i^{\text{ème}}$ mesure donne

$$(6) \quad Y_i = a_1 M_{1i} + a_2 M_{2i} + \dots + a_m M_{mi}.$$

L'erreur de Y_i sera

$$\bar{\varepsilon}_i = Y_i - Y = a_1 (M_{1i} - z_1) + \dots + a_m (M_{mi} - z_m) = a_1 \varepsilon_{1i} + \dots + a_m \varepsilon_{mi}.$$

On tire de l'équation (6) la moyenne des Y_i

$$\frac{\sum Y_i}{n} = a_1 \frac{\sum M_{1i}}{n} + \dots + a_m \frac{\sum M_{mi}}{n}.$$

De plus l'écart $\bar{\varepsilon}_i$ entre Y_i et sa moyenne sera

$$\bar{\varepsilon}_i = Y_i - \frac{\sum Y_i}{n} = a_1 \left[M_{1i} - \frac{\sum M_{1i}}{n} \right] + \dots + a_m \left[M_{mi} - \frac{\sum M_{mi}}{n} \right]$$

ou encore

$$\bar{\xi}_i = a_1 \xi_{1i} + \dots + a_m \xi_{mi},$$

où les ξ_{vi} sont les écarts respectifs.

Enfin $\bar{\eta}$, l'erreur de la moyenne des Y_i sera

$$(7) \quad \bar{\eta} = \frac{\sum Y_i}{n} - Y = a_1 \left[\frac{\sum M_{1i}}{n} - z_1 \right] + \dots + a_m \left[\frac{\sum M_{mi}}{n} - z_m \right],$$

on trouve donc comme précédemment $\bar{\varepsilon}_i = \bar{\xi}_i + \bar{\eta}$.

La probabilité d'obtenir les erreurs $\varepsilon_{1i}, \dots, \varepsilon_{mi}$ est tirée de la formule (1)

$$(8) \quad \frac{1}{\zeta_1 \dots \zeta_m (2\pi)^{\frac{1}{2}m}} e^{-\frac{\sum \varepsilon_{vi}^2}{2\zeta_v^2}} \Delta \varepsilon_1 \dots \Delta \varepsilon_m.$$

Pour obtenir la probabilité $\mathfrak{P}(\lambda)$ que la valeur absolue de l'erreur de la moyenne des Y_i soit plus petite que λ , il suffit de faire la somme de la probabilité (8) pour toutes les valeurs de ε_{vi} telles que

$$|\bar{\eta}| = |a_{1i} \varepsilon_{1i} + \dots + a_{mi} \varepsilon_{mi}| < \lambda.$$

On y arrive le plus simplement si on multiplie (8) par le facteur discontinu $F(\lambda, \sum a_v^2 \varepsilon_{vi})$ qui est égal à l'unité pour les valeurs satisfaisant à l'inégalité précédente et nul pour les autres; alors on peut effectuer les sommations par rapport aux ε_v de $-\infty$ à ∞ ; le même calcul que celui qu'on a employé pour déterminer la probabilité de l'erreur de la moyenne au § 4, conduit à

$$\mathfrak{P}(\lambda) = \frac{2}{\pi} \int_0^\infty e^{-\frac{1}{2}u^2} \sin u\lambda \frac{du}{u}.$$

Posons maintenant $\omega = u \sqrt{\sum a_v^2 \zeta_v^2}$ et $u\lambda = \omega x$; on obtient

$$\mathfrak{P}(\lambda) = \frac{2}{\pi} \int_0^\infty e^{-\frac{1}{2}\omega^2} \sin \omega x \frac{d\omega}{\omega},$$

où $x = \lambda / \sqrt{\sum a_v^2 \zeta_v^2} = \lambda / \bar{\zeta}$.

Par suite

$$(9) \quad \bar{\zeta}^2 = \sum a_v^2 \zeta_v^2$$

est la moyenne des carrés des erreurs des Y_i . On en conclut que si les z_v suivent la loi (1), la fonction Y la suivra aussi, de plus l'écart quadratique de Y sera $\bar{\zeta}$.

Cas particulier. Soit $Y = \frac{1}{m}(z_1 + z_2 + \dots + z_m)$, autrement dit Y est la moyenne des quantités z_v ; en outre si l'on a $\zeta_v = \zeta$, alors de (9) il résulte que l'erreur quadratique de la moyenne des z_v sera

$$\bar{\zeta} = \sqrt{m \frac{1}{m^2} \zeta^2} = \frac{\zeta}{\sqrt{m}}$$

conformément à ce qu'on a trouvé précédemment, mais cette fois m n'est pas nécessairement grand.

§ 6. Etant données une grandeur z et son erreur quadratique déterminée préalablement, on fait n mesures M_1, M_2, \dots, M_n et l'on demande la *probabilité pour que l'erreur quadratique correspondant aux observations soit plus petite que λ* . Comme $M_i - z = \varepsilon_i$ cette erreur quadratique est $\sqrt{\sum \varepsilon_i^2/n}$; par suite, il faut déterminer la probabilité d'avoir

$$(10) \quad \sum \varepsilon_i^2 < n\lambda^2.$$

La probabilité d'obtenir les mesures M_i , ou ce qui revient au même, la probabilité des erreurs ε_i est donnée par

$$(11) \quad \left(\frac{1}{\zeta \sqrt{2\pi}} \right)^n e^{-\sum \varepsilon_i^2 / 2\zeta^2} d\varepsilon_1 \dots d\varepsilon_n.$$

Pour obtenir la probabilité cherchée il faut faire la somme des probabilités (11) correspondant à chaque système ε_i satisfaisant à l'inégalité (10). On détermine cette somme, en faisant correspondre à chaque système de ε_i un point de l'espace à n dimensions. Les points favorables seront à l'intérieur de la sphère de rayon $R = \lambda\sqrt{n}$. La probabilité est constante à l'intérieur de la couche sphérique de rayon r et $r + dr$, où

$$r^2 = \varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_n^2.$$

Le volume de cette couche est égal à

$$\frac{2\pi^{\frac{1}{2}n} r^{n-1}}{\Gamma\left(\frac{1}{2}n\right)} dr.$$

La probabilité cherchée est proportionnelle à ce volume et à la probabilité (11). On a donc

$$(12) \quad d\mathfrak{P} = \frac{2C}{\Gamma\left(\frac{1}{2}n\right)} \left(\frac{1}{\zeta\sqrt{2}} \right)^n r^{n-1} e^{-r^2/2\zeta^2} dr.$$

Pour déterminer la valeur de C remarquons que l'intégrale de cette expression, r variant de zéro à ∞ , doit être égale à l'unité; en effet c'est la probabilité pour que l'erreur quadratique soit quelconque. Pour effectuer l'intégration introduisons une nouvelle variable $\omega = r^2/2\zeta^2$ d'où $d\omega = r dr/\zeta^2$. On a

$$\int_0^{\infty} d\mathfrak{P} = \frac{C}{\Gamma\left(\frac{1}{2}n\right)} \int_0^{\infty} \omega^{\frac{1}{2}n-1} e^{-\omega} d\omega = C = 1.$$

Par suite dans la formule (12), la constante C est égale à l'unité.

Finalement dans le cas de la probabilité cherchée on doit avoir $r^2 < n\lambda^2$; par suite $\omega < n\lambda^2/2\zeta^2 = x$; et

$$(13) \quad \mathfrak{P}(n\lambda^2) = \frac{1}{\Gamma\left(\frac{1}{2}n\right)} \int_0^x \omega^{\frac{1}{2}n-1} e^{-\omega} d\omega = I(\bar{u}, \bar{p})$$

où $I(\bar{u}, \bar{p})$ est le rapport de la fonction gamma incomplète $\Gamma_x\left(\frac{1}{2}n\right)$ à la fonction complète correspondante $\Gamma\left(\frac{1}{2}n\right)$. On a

$$\bar{p} = \frac{1}{2}n - 1 \quad \text{et} \quad \bar{u} = \frac{\lambda^2}{\zeta^2} \sqrt{\frac{1}{2}n}.$$

Cas particulier 1. La probabilité pour que l'on ait $\Sigma \varepsilon_i^2 < n\zeta^2$ est donnée par la formule (13)

$$\mathfrak{P}(n\zeta^2) = I\left(\sqrt{\frac{1}{2}n}, \frac{1}{2}n - 1\right).$$

Cette probabilité est d'après les tables de la fonction gamma incomplète un peu plus grand qu'une demi.

Cas particulier 2. En posant $n=1$ dans la formule (13) on obtient la probabilité pour que le carré de l'erreur d'une observation soit plus petit que λ^2 . On trouve

$$\frac{1}{\sqrt{\pi}} \int_0^{\lambda^2/2\zeta^2} \omega^{-\frac{1}{2}} e^{-\omega} d\omega = I\left(\frac{\lambda^2}{\zeta^2\sqrt{2}}, -\frac{1}{2}\right),$$

c'est aussi la probabilité pour que la valeur absolue de l'erreur d'une mesure soit plus petite que λ ; elle est donnée par la formule de LAPLACE:

$$\frac{2}{\sqrt{2\pi}} \int_0^{\lambda/\zeta} e^{-\frac{1}{2}t^2} dt.$$

On peut donc exprimer ce résultat par une fonction gamma-incomplète

$$I\left(\frac{\lambda^2}{\zeta^2\sqrt{2}}, -\frac{1}{2}\right) = \frac{2}{\sqrt{2\pi}} \int_0^{\lambda/\zeta} e^{-t^2/2} dt.$$

Remarque. L'intégration par parties de $I(\bar{u}, \bar{p})$ dans (13) donne lorsque $n = 2\nu$

$$I(\bar{u}, \bar{p}) = 1 - e^{-x} \sum_{i=0}^{\nu-1} \frac{x^{\nu-1-i}}{\Gamma(\nu-i)}$$

et lorsque $n = 2\nu + 1$

$$I(\bar{u}, \bar{p}) = I\left(x\sqrt{2}, -\frac{1}{2}\right) - e^{-x} \sum_{i=0}^{\nu-1} \frac{x^{\nu-\frac{1}{2}-i}}{\Gamma\left(\nu+\frac{1}{2}-i\right)}.$$

Dans ces formules $\bar{u} = x\sqrt{2}/\sqrt{n}$ et $\bar{p} = \frac{1}{2}n - 1$.

Pour obtenir la probabilité que la *moyenne des carrés des erreurs* soit comprise entre λ^2 et $\lambda^2 + d\lambda^2$ il suffit de poser dans (12) $r^2 = n\lambda^2$; on obtient

$$(14) \quad \frac{n}{2\zeta^2 \Gamma\left(\frac{1}{2}n\right)} \left(\frac{n\lambda^2}{2\zeta^2}\right)^{\frac{1}{2}n-1} e^{-n\lambda^2/2\zeta^2} d\lambda^2.$$

Cette probabilité est nulle pour $\lambda = 0$ et $\lambda = \infty$, elle est maximum pour $\lambda^2 = \frac{n-2}{n} \zeta^2$.

En multipliant la formule (14) par λ^2 et en intégrant de zéro à ∞ , on a l'espérance mathématique $\mathcal{E}(\lambda^2)$ des grandeurs λ^2 . La méthode appliquée dans le cas de la formule (12) conduit à $\mathcal{E}(\lambda^2) = \zeta^2$.

En multipliant la formule (14) par λ^4 et en intégrant on a l'espérance mathématique des λ^4 , $\mathcal{E}(\lambda^4) = \frac{n+2}{n} \zeta^4$.

Finalement le carré de l'erreur quadratique de λ^2 sera

$$\mathcal{E}(\lambda^2 - \zeta^2)^2 = \mathcal{E}(\lambda^4) - \mathcal{E}^2(\lambda^2) = \frac{2}{n} \zeta^4;$$

enfin l'erreur quadratique de λ^2 est $\zeta^2 \sqrt{\frac{2}{n}}$.

§ 7. Si z était connu et ζ inconnu on admettra que la probabilité à priori de ζ est proportionnelle à $\Delta\zeta/\zeta^\alpha$. Après avoir fait n mesures, $\Delta V_1(\zeta)$, la probabilité à posteriori de ζ sera donnée par le théorème de BAYES. En posant dans (3)

$$s^2 = \Sigma \varepsilon_i^2 = n\eta^2 + \Sigma \xi_i^2 = n(\eta^2 + \sigma^2)$$

on aura

$$\Delta V_1(\zeta) = \frac{\zeta^{-n-\alpha} e^{-s^2/2\zeta^2} \Delta\zeta}{\int_0^\infty \zeta^{-n-\alpha} e^{-s^2/2\zeta^2} d\zeta}.$$

Transformons cette intégrale en posant $\omega = s^2/2\zeta^2$ et par suite $d\omega = -s^2 d\zeta/\zeta^3$; le dénominateur devient :

$$(15) \quad \frac{2^{\frac{1}{2}(n+\alpha-3)}}{s^{n-1+\alpha}} \int_0^\infty \omega^{\frac{1}{2}(n+\alpha-3)} e^{-\omega} d\omega = \frac{2^{\frac{1}{2}(n+\alpha-3)}}{s^{n+\alpha-1}} \Gamma\left(\frac{n+\alpha-1}{2}\right).$$

Enfin la probabilité à posteriori de ζ sera si z est connu :

$$\Delta V_1(\zeta) = \frac{\omega^{\frac{1}{2}(n+\alpha-3)} e^{-\omega} \Delta\omega}{\Gamma\left(\frac{n+\alpha-1}{2}\right)} = \frac{s^{n+\alpha-1} e^{-s^2/2\zeta^2} \Delta\zeta}{2^{\frac{1}{2}(n+\alpha-3)} \Gamma\left(\frac{n+\alpha-1}{2}\right) \zeta^{n+\alpha}}.$$

Cette probabilité sera maximum pour $\zeta^2 = \frac{n}{n+\alpha} (\eta^2 + \sigma^2) = \frac{\Sigma \varepsilon_i^2}{n+\alpha}$.

§ 8. Lorsque z est inconnu, mais que la précision ζ est connue, ayant été déterminée par des observations antérieures, alors la probabilité (1) sera maximum quand la somme des carrés des erreurs $\Sigma \varepsilon_i^2 = \Sigma (M_i - z)^2$ est minimum. On en conclut que la loi d'ADRAIN (1) conduit aussi au principe des moindres carrés.

Lorsque z et ζ sont inconnus et que l'on admet que leur probabilités à priori sont respectivement proportionnelles à Δz et $\Delta\zeta/\zeta^\alpha$ alors, après les n mesures, la probabilité à posteriori de z et ζ sera

$$(16) \quad \Delta V_2(z, \zeta) = \frac{\zeta^{-n-\alpha} e^{-\Sigma (M_i - z)^2/2\zeta^2} \Delta z \Delta\zeta}{\int_{-\infty}^\infty \int_0^\infty \zeta^{-n-\alpha} e^{-\Sigma (M_i - z)^2/2\zeta^2} dz d\zeta}.$$

Pour obtenir la probabilité de z , on remplace quelquefois dans cette formule ζ par sa valeur la plus probable, mais ce procédé ne conduit pas à des résultats satisfaisants. Il vaut mieux déterminer la probabilité totale de z , quel que soit ζ ; il suffit

pour cela d'intégrer le numérateur et le dénominateur par rapport à ξ . En vertu de (15) l'intégration du numérateur donnera (comme $\Delta z = \Delta \eta$)

$$\frac{2^{\frac{1}{2}(n+\alpha-3)} \Gamma\left(\frac{n+\alpha-1}{2}\right)}{[n(\sigma^2 + \eta^2)]^{\frac{1}{2}(n+\alpha-1)}} \Delta \eta;$$

il en résulte

$$\Delta V(z) = \frac{(\sigma^2 + \eta^2)^{-\frac{1}{2}(n+\alpha-1)} \Delta \eta}{\int_0^\infty (\sigma^2 + \eta^2)^{-\frac{1}{2}(n+\alpha-1)} d\eta}.$$

Posons $\tan \varphi = \eta/\sigma$ et par suite $d\eta = \sigma d\varphi / \cos^2 \varphi$. L'intégrale du dénominateur sera :

$$\sigma^{-n-\alpha+2} \int_0^{\frac{1}{2}\pi} \cos^{n+\alpha-3} \varphi d\varphi = \frac{\pi \Gamma(n+\alpha-2)}{(2\sigma)^{n+\alpha-2} \Gamma^2\left(\frac{n+\alpha-1}{2}\right)}.$$

(Voir BIERENS DE HAAN, Table 41, formule 3). Par suite la probabilité à posteriori de z et en même temps la probabilité à posteriori de l'erreur η de la moyenne des mesures sera quelle que soit la précision inconnue ξ

$$\Delta V(z) = \frac{\Gamma^2\left(\frac{n+\alpha-1}{2}\right)}{\Gamma(n+\alpha-2)} \cdot \frac{1}{2\pi\sigma} \left(\frac{4\sigma^2}{\sigma^2 + \eta^2}\right)^{\frac{1}{2}(n+\alpha-1)} \Delta \eta$$

ou encore

$$(17) \quad \Delta V(z) = \frac{2^{n+\alpha-2} \Gamma^2\left(\frac{n+\alpha-1}{2}\right)}{\pi \Gamma(n+\alpha-2)} \cos^{n+\alpha-3} \varphi d\varphi.$$

Cette probabilité est maximum lorsque $\eta = 0$ ou $z = \Sigma M_i/n$; mais alors z rend minimum la somme des carrés des erreurs; ainsi le principe des moindres carrés est valable dans ce cas aussi.

De la formule (17) il suit immédiatement la probabilité \mathfrak{S} pour que la valeur absolue de l'erreur de la moyenne soit plus petite que λ , c'est-à-dire $|\eta| < \lambda$. Comme $\eta = \sigma \tan \varphi$ et $\lambda = \sigma \tan \varphi_1$, il faut avoir $\varphi < \varphi_1$, où $\varphi_1 = \arctan \lambda/\sigma$. On trouve quelle que soit la précision ξ

$$\mathfrak{S} = \frac{2^{n+\alpha-2} \Gamma^2\left(\frac{n+\alpha-1}{2}\right)}{\pi \Gamma(n+\alpha-2)} \int_0^{\varphi_1} \cos^{n+\alpha-3} \varphi d\varphi.$$

Exemple 1. $n=4$, $\alpha=0$, on a

$$\mathcal{S} = \sin \varphi_1 = \sqrt{\frac{\lambda^2}{\sigma^2 + \lambda^2}}$$

et si $\lambda = \sigma$ alors $\mathcal{S} = 1/\sqrt{2}$.

Pour obtenir des valeurs acceptables depuis $n=1$ il faut poser au moins $\alpha=2$, alors on trouve pour $\lambda = \sigma$:

$$n=1 \\ \mathcal{S} = \frac{1}{2}$$

$$n=2 \\ \mathcal{S} = 1/\sqrt{2}$$

$$n=3 \\ \mathcal{S} = \frac{1}{\pi} + \frac{1}{2}.$$

De la même manière on peut déterminer la probabilité de ζ quel que soit z , en intégrant le numérateur et le dénominateur de (16) par rapport à z . Le numérateur sera :

$$\zeta^{-n-\alpha} e^{-n\sigma^2/2\zeta^2} \Delta \zeta \int_{-\infty}^{\infty} e^{-n\eta^2/2\zeta^2} d\eta,$$

en posant $t = \eta\sqrt{n}/\zeta$ et par suite $d\eta = \zeta dt/\sqrt{n}$, on obtient

$$\frac{\sqrt{2\pi} e^{-n\sigma^2/2\zeta^2} \Delta \zeta}{\zeta^{n+\alpha-1} \sqrt{n}}$$

Comme le dénominateur donne une valeur semblable on aura

$$V(\zeta) = \frac{\zeta^{-n-\alpha+1} e^{-n\sigma^2/2\zeta^2} \Delta \zeta}{\int_0^\infty \zeta^{-n-\alpha+1} e^{-n\sigma^2/2\zeta^2} d\zeta}.$$

Or l'intégrale du dénominateur est d'après (15) égale à

$$2^{\frac{1}{2}(n+\alpha-2)} \Gamma\left(\frac{n+\alpha-1}{2}\right) / (n\sigma^2)^{\frac{1}{2}(n+\alpha-2)};$$

il s'ensuit la probabilité à posteriori de ζ quelle que soit la valeur inconnue de z :

$$V(\zeta) = \frac{(n\sigma^2)^{\frac{1}{2}(n+\alpha-2)} e^{-n\sigma^2/2\zeta^2}}{2^{\frac{1}{2}(n+\alpha-2)} \Gamma\left(\frac{n+\alpha-1}{2}\right) \zeta^{n+\alpha-1}} \Delta \zeta.$$

Le maximum de cette probabilité a lieu pour $\zeta^2 = \frac{n\sigma^2}{n+\alpha-1}$

En posant dans cette formule $\alpha=0$, la valeur la plus probable de ζ^2 serait $\Sigma e_i^2/(n-1)$ conformément au résultat proposé par GAUSS.

§ 9. Dans le problème du § 5 les coefficients a_1, a_2, \dots, a_m étaient donnés et on a mesuré les grandeurs z_1, z_2, \dots, z_m , puis on a cherché à déterminer les probabilités de $Y = a_1 z_1 + a_2 z_2 + \dots + a_m z_m$. Il y a une autre catégorie de problèmes. Étant donnés les coefficients a_{i1}, \dots, a_{im} pour les valeurs de $i = 1, 2, \dots, n$, on mesure les grandeurs suivantes

$$Y_i = \sum_{v=1}^m a_{iv} x_v.$$

Les quantités x_1, x_2, \dots, x_m sont inconnues et on demande leur probabilité. Soit M_i la mesure de Y_i , l'erreur inconnue sera :

$$\varepsilon_i = M_i - Y_i.$$

En supposant que les Y_i sont du même ordre de grandeur, et que la précision des observations est la même pour toutes les mesures l'erreur quadratique inconnue étant ζ , la probabilité à posteriori des grandeurs Y_i et par suite aussi des erreurs ε_i sera :

$$\Delta P_1(\varepsilon_1, \dots, \varepsilon_n, \zeta) = \frac{\zeta^{-n-\alpha} e^{-\sum \varepsilon_i^2 / 2 \zeta^2} \Delta \varepsilon_1 \dots \Delta \varepsilon_n \Delta \zeta}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_0^{\infty} \zeta^{-n-\alpha} e^{-\sum \varepsilon_i^2 / 2 \zeta^2} d\varepsilon_1 \dots d\varepsilon_n d\zeta}.$$

En effectuant l'intégration au numérateur et au dénominateur par rapport à ζ comme dans le cas de la formule (16) on a la probabilité à posteriori des Y_i quelle que soit la précision ζ

$$(18) \quad \Delta P(\varepsilon_1, \dots, \varepsilon_n) = \frac{(\sum \varepsilon_i^2)^{-\frac{1}{2}(n+\alpha-1)} \Delta \varepsilon_1 \dots \Delta \varepsilon_n}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\sum \varepsilon_i^2)^{-\frac{1}{2}(n+\alpha-1)} d\varepsilon_1 \dots d\varepsilon_n}.$$

Cette probabilité sera encore maximum lorsque $\sum \varepsilon_i^2$ est minimum. Il s'agit donc de rendre minimum la quantité :

$$(19) \quad \sum \varepsilon_i^2 = \sum_i (M_i - a_{i1}x_1 - a_{i2}x_2 - \dots - a_{im}x_m)^2.$$

Les équations de condition du minimum, dites équations normales, seront :

$$(20) \quad \frac{\partial \sum \varepsilon_i^2}{\partial x_v} = \sum_i a_{iv} (M_i - a_{i1}x_1 - a_{i2}x_2 - \dots - a_{im}x_m) = 0$$

pour $v = 1, 2, \dots, m$. La somme des carrés des erreurs minima est

$$(21) \quad \sum_i M_i (M_i - a_{i1}x_1 - a_{i2}x_2 - \dots - a_{im}x_m) = \omega^2;$$

en effet en multipliant les équations (20) par x_v et en retranchant

les produits de l'équation précédente, on obtient l'équation (19). La valeur minimum ω^2 de $\Sigma \varepsilon_i^2$ est importante, car la probabilité pour que $\Sigma \varepsilon_i^2$ soit plus petite que ce minimum est nulle. Il faudra en tenir compte en effectuant les intégrations dans la formule (18).

Pour déterminer ω posons :

$$M_i = -a_{i, m+1}.$$

et

$$x_{m+1} = 1;$$

de plus

$$\sum_{i=1}^n a_{i\nu} a_{is} = b_{\nu s}$$

donc

$$b_{\nu s} = b_{s\nu}.$$

Soit

$$D_i = \begin{vmatrix} b_{11} & b_{12} & \dots & b_{1i} \\ b_{21} & b_{22} & \dots & b_{2i} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \dots & b_{ii} \end{vmatrix},$$

alors les m équations normales s'écriront

$$b_{1\nu}x_1 + b_{2\nu}x_2 + \dots + b_{m+1,\nu}x_{m+1} = 0$$

pour $\nu = 1, 2, \dots, m$.

En outre l'équation (21) devient :

$$b_{1,m+1}x_1 + b_{2,m+1}x_2 + \dots + b_{m+1,m+1}x_{m+1} = \omega^2.$$

La résolution de ces $m+1$ équations à $m+1$ inconnues x_1, \dots, x_{m+1} conduit à

$$x_{m+1} = \omega^2 D_m / D_{m+1};$$

mais comme $x_{m+1} = 1$, on a $\omega^2 = D_{m+1} / D_m$.

Les premières m équations donnent, si l'on pose $x_{m+1} = 1$,

$$(22) \quad x_\nu = \frac{1}{D_m} \sum_{s=1}^m (-1)^{\nu+s+1} b_{m+1,s} \beta_{\nu s}$$

où $\beta_{\nu s}$ est le mineur de D_m obtenu en supprimant la ligne ν et la colonne s dans ce déterminant.

En partant de (22) on peut déterminer l'erreur quadratique ζ_ν de x_ν lorsqu'on connaît l'erreur quadratique ξ des mesures M . En effet en remplaçant $b_{m+1,s}$ par sa valeur obtenue ci-dessus on trouve

$$x_\nu = \sum_{i=1}^n M_i \sum_{s=1}^m \frac{(-1)^{s+\nu}}{D_m} \beta_{\nu s} a_{is}$$

et en vertu de (9) le carré de l'erreur quadratique (ou la dispersion) de x_ν sera

$$\zeta_\nu^2 = \zeta^2 \sum_i \sum_s \sum_k \frac{(-1)^{s+k}}{D_m^2} \beta_{\nu s} \beta_{\nu k} a_{is} a_{ik};$$

mais

$$\sum_k (-1)^{s+k} \beta_{\nu k} b_{sk}$$

est égal à zéro si $s \neq \nu$ et égal à D_m si $s = \nu$; on en conclut que le carré de l'erreur quadratique de x_ν est

$$\zeta_\nu^2 = \frac{\beta_{\nu \nu}}{D_m} \zeta^2.$$

Lorsqu'on attribue aux mesures M_i l'unité de *poids*, alors le poids de la valeur x_ν obtenue (22), étant inversement proportionnelle aux dispersions respectives, sera

$$\frac{\zeta^2}{\zeta_\nu^2} = \frac{D_m}{\beta_{\nu \nu}}.$$

Pour intégrer l'équation (18) remarquons de nouveau que $\Sigma \varepsilon_i^2 = r^2$ est constant sur la surface de la sphère de rayon r dans l'espace à n dimensions; cette surface ayant pour aire $2\pi^{\frac{1}{2}n} r^{n-1} / \Gamma\left(\frac{n}{2}\right)$, donc l'intégrale du dénominateur de (18) devient

$$\int_{\omega}^{\infty} \frac{2\pi^{\frac{1}{2}n}}{\Gamma\left(\frac{n}{2}\right)} \frac{dr}{r^\alpha} = \frac{2\pi^{\frac{1}{2}n}}{\Gamma\left(\frac{n}{2}\right)} \frac{1}{(\alpha-1)\omega^{\alpha-1}},$$

il en résulte que la probabilité (18) est égale à

$$\begin{aligned} (23) \quad \Delta P(\varepsilon_1, \dots, \varepsilon_n) &= \\ &= \frac{(\alpha-1) \Gamma\left(\frac{n}{2}\right)}{2\pi^{\frac{1}{2}n}} \left(\frac{D_{m+1}}{D_m}\right)^{\frac{1}{2}(\alpha-1)} (\Sigma \varepsilon_i^2)^{-\frac{1}{2}(n+\alpha-1)} \Delta \varepsilon_1 \dots \Delta \varepsilon_n \end{aligned}$$

et si l'on transforme le numérateur de (18) de la même manière que le dénominateur on trouve

$$\Delta P(\varepsilon_1, \dots, \varepsilon_n) = (\alpha-1) \left(\frac{D_{m+1}}{D_m}\right)^{\frac{1}{2}(\alpha-1)} \frac{dr}{r^\alpha}.$$

Cette formule donne immédiatement la probabilité pour que la somme des carrés des erreurs des Y_i soit plus petite que λ^2 c'est-à-dire $\sum \varepsilon_i^2 = r^2 < \lambda^2$, quelle que soit la précision inconnue ζ . Il suffit d'intégrer l'expression précédente de zéro à λ , on trouve

$$\mathcal{P}(\lambda^2) = 1 - \left(\frac{D_{m+1}}{\lambda^2 D_m} \right)^{\frac{1}{2}(\alpha-1)}$$

Ce résultat a été obtenu en supposant que la probabilité à priori de la précision ζ était proportionnelle à $\Delta\zeta/\zeta^\alpha$. Remarquons que pour $\alpha=0$ (GAUSS) la probabilité $\mathcal{P}(\lambda^2)$ diminuerait si λ^2 augmente, ce qui est impossible; de plus pour $\alpha=1$ on aurait $\mathcal{P}(\lambda^2)=1$ quel que soit λ , c'est inadmissible aussi. On en conclut comme cela a été remarqué précédemment que dans la probabilité à priori de ζ il faut poser au moins $\alpha=2$.

Lorsqu'on a mesuré n fois une grandeur inconnue, ne connaissant pas la précision ζ^2 des mesures, mais sachant que sa valeur la plus probable est (voir la fin du § 8)

$$\frac{\sum \varepsilon_i^2}{n + \alpha - 1} = \frac{n\sigma^2}{n + \alpha - 1},$$

on trouve dans l'hypothèse de GAUSS correspondant à $\alpha=0$ que cette valeur sera supérieure à σ^2 qui est le carré de l'écart quadratique des mesures. Dans l'hypothèse $\alpha=1$ elle sera égale à σ^2 et enfin dans l'hypothèse $\alpha \geq 2$, elle sera inférieure à σ^2 . Nous avons vu que c'est cette dernière hypothèse qui est la mieux justifiée. Il en résulte que si l'on se contente d'une approximation pour la valeur la plus probable de la précision ζ^2 , on peut prendre $\zeta^2 = \sigma^2$; mais comme c'est déjà plus grande que la valeur rigoureuse, on n'a aucune raison d'employer celle proposée par GAUSS, qui est encore plus grande.

(Reçu le 18 octobre 1941)